PATENT
Attorney Docket No. CIT1100
(06618/337001)

X NEW PATENT APPLICATION
___ CONTINUATION-IN-PART
___ DIVISIONAL

CERTIFICATE OF MAILING BY "EXPRESS MAIL"

"EXPRESS MAIL" MAILING LABEL NO EL360646174US
DATE OF DEPOSIT August 3, 1999 I HEREBY CERTIFY THAT THIS PAPER OR FEE IS BEING DEPOSITED WITH THE UNITED STATES POSTAL SERVICE "EXPRESS MAIL POST OFFICE TO ADDRESSEE" SERVICE UNDER 37 C.F.R. 1.10 ON THE DATE INDICATED ABOVE AND IS ADDRESSED TO THE ASSISTANT COMMISSIONER FOR PATENTS, WASHINGTON, D.C. 20231

ANGIE VARGAS
(TYPED OR PRINTED NAME OF PERSON MAILING PAPER OR FEE)

SIGNATURE OF PERSON MAILING (PAPER OR FEE)

ASSISTANT COMMISSIONER
FOR PATENTS
Box Patent Application
Washington, D.C. 20231

Sir:

Transmitted herewith for filing is the provisional-to-utility patent application of

Inventor: WILLIAM J. DREYER

For: **METHOD OF DETERMINING CELL OR TISSUE TYPE BY TRANSMEMBRANE RECEPTOR IDENTIFIERS**

This is a request for filing a new utility patent application relying on provisional application Serial No. 60/095,148, filed August 3, 1998, now pending.

Enclosed are:

X 51 pages of the Specification, which includes 9 pages of the claims, plus 1 page of the Abstract;

X 2 sheets of drawing(s) ___ Formal; X Informal;

X An assignment of the invention to California Institute of Technology ;

X A Declaration, unsigned, and

___ A verified statement to establish small entity status under 37 C.F.R. 1.9 and 37 C.F.R. 1.27.

FULL NAME OF FIRST INVENTOR	*LAST NAME:*	*FIRST NAME:*	*MIDDLE NAME:*
	DREYER	WILLIAM	J.
CITIZENSHIP	*STATE OR FOREIGN COUNTRY:* USA		
POST OFFICE ADDRESS	*POST OFFICE ADDRESS:*	*CITY AND STATE:*	*ZIP CODE:*
Pasadena, CA 91103	1875 Devon Road	Pasadena, CA	91103

The filing fee has been calculated as shown below:

For	Number Filed		Number Extra		Rate			Fee	
					Small Entity	Other Entity		Small Entity	Other Entity
Total Claims	49	-20 =	29	X	$09	$18	=	261.00	
Independent Claims	8	- 3 =	5	X	$39	$78	=	195.00	
Multiple Dependent Claims Presented: ___ Yes _X_ No					$130	$260	=		
				BASIC FEE	$380	$760	=	380.00	
					TOTAL FEE			836.00	

X The payment of the filing fee is to be deferred until the executed Declaration is filed. Do not charge our deposit account.

___ The Assistant Commissioner is hereby authorized to charge payment of the following fees associated with this communication or credit any overpayment to Deposit Account No. 07-1895. A duplicate copy of this sheet is enclosed.

___ Any additional filing fees required under 37 C.F.R. 1.16.
___ Any patent application processing fees under 37 C.F.R. 1.17.

Respectfully submitted,

Date: 8/3/99

Lisa A. Haile, Ph.D.
Attorney for Applicant
Registration No. 38,347
Telephone: 858-677-1456
Facsimile: 858-677-1465

GRAY CARY WARE & FREIDENRICH LLP
4365 Executive Drive, Suite 1600
San Diego, CA 92121-2189

CERTIFICATE OF MAILING BY "EXPRESS MAIL"

"EXPRESS MAIL LABEL NUMBER EL340646174US

DATE OF DEPOSIT August 3, 1999

I HEREBY CERTIFY THAT THIS PAPER OR FEE IS BEING DEPOSITED WITH THE UNITED STATES POSTAL SERVICE "EXPRESS MAIL POST OFFICE TO ADDRESSEE" SERVICE UNDER 37 CFR 1 10 ON THE DATE INDICATED ABOVE AND IS ADDRESSED TO THE ASSISTANT COMMISSIONER OF PATENTS, BOX PATENT APPLICATION, WASHINGTON, D C 20231

ANGIE VARGAS
(TYPED OR PRINTED NAME OF PERSON MAILING PAPER)

(SIGNATURE OF PERSON MAILING PAPER OR FEE)

APPLICATION

for

UNITED STATES LETTERS PATENT

on

METHOD OF DETERMINING CELL OR TISSUE TYPE BY TRANSMEMBRANE RECEPTOR IDENTIFIERS

by

WILLIAM J. DREYER

Sheets of Drawings: Two (2)

Docket No.: CIT1100

Attorneys

Gray Cary Ware & Freidenrich
4365 Executive Drive, Suite 1600
San Diego, California 92121-2189

METHOD OF DETERMINING CELL OR TISSUE TYPE BY TRANSMEMBRANE RECEPTOR IDENTIFIERS

REFERENCE TO RELATED APPLICATIONS

The present application is claims priority under 35 U.S.C. 119(e)(1) to U.S. Provisional Patent application Serial Number 60/095,148, filed August 3, 1998, which is incorporated by reference in its entirety.

FIELD OF THE INVENTION

The present invention relates generally to cell lineages, organogenesis and embryogenesis, and more specifically to methods for identifying and characterizing cell and tissue types based on the surface display of molecules, such as olfactory ("serpentine") receptors, and/or based on the DNA "fingerprint" of a cell.

BACKGROUND OF THE INVENTION

The "Area Code Hypothesis" helps explain how chromosomes sculpture living organisms. The DNA contained in the two cells that will form identical twins is able to choreograph the parallel development of two strikingly similar individuals through birth and through all of the stages of their lives. In a favorable environment the twins will grow, rearrange their bodies at puberty, and go through the changes of maturity and aging in parallel. Even the MRI images of their brains will be strikingly similar and very different from other brain images. It was consideration of this extraordinary precision of cell and neural assembly that originally lead to the proposition of the Area Code Hypothesis (1). The hypothesis was based on extensive genetic, molecular, and cellular studies of the immune system (2,3; see also refs. in (1)).

Key elements of the hypothesis are the following: 1) Large multigene families must exist that code for cell surface receptors providing highly specific cell-cell recognition functions; 2) Receptors must be used repeatedly in a combinatorial fashion so that a finite number of genes can provide enough information to generate the required large number of cellular addresses; 3) Programmed genetic switching similar in some respects to that seen during the development of the immune system is assumed to aid in the complex control of the expression of these address codes in specific lineages and cells (4); and 4) Some classes of cell surface receptors are assumed to be widely expressed throughout the organism and code for large regions resembling, for example, the country codes of our telephone dialing system. Other classes of molecules would be more restricted in expression and are expected to code for multiple smaller regions of the embryo somewhat comparable, according to this metaphor, to the multiple regions specified by area codes and regional prefixes throughout the world. Finally, it is assumed that molecules exist that encode a specific cellular address comparable to the four digits used to code for a single, specific telephone in any one of the numerous, distinct topological regions specified by the earlier codes. Both the telephone digits and the genes and cell surface receptors that provide this last part of the code may be used repeatedly in diverse physical locations.

SUMMARY OF THE INVENTION

The present invention is based on the seminal discovery that cell surface display molecules, such as certain transmembrane receptors, on cells can be used to identify cell lineages. Detection of such receptors gives one the ability to manipulate the decisions that cells make and allows one of skill in the art to isolate specific cell lineages or manipulate the lineage decision made by a particular cell.

One embodiment of the invention provides a method of obtaining a specific cell type or lineage. The method includes obtaining a sample of cells, contacting the cells with an agent, such as an antibody or a ligand specific for a cell surface marker such that the antibody or ligand binds to a cell in the sample and separating the cell that is bound by the antibody or ligand from the sample thereby obtaining a population of a specific cell type or lineage. The cell population may be further purified by selecting for cells by expression of at least one additional marker associated with a specific cell type. The method of the invention includes identifying a cell type by detecting expression of at least one serpentine receptor, wherein the presence or absence of the serpentine receptor is indicative of a cell type. The selection of this subpopulation of cells may be done by positive or negative selection methods.

In another embodiment the invention provides a method of identifying a cell type by detecting expression of an serpentine receptor specific for a certain cell or tissue type. Identification may be through detecting expression of serpentine polypeptides encoding serpentine polypeptides. Exemplary serpentine polypeptides include olfactory receptors (Dreyer, W., *Proc Natl. Acad. Sci.*, 95: 9072-9077, 1998, herein incorporated by reference).

In a further embodiment the invention provides a method of treating a cell proliferative disorder in a subject. The method includes administering an effective amount of an antibody coupled to a first factor wherein the antibody binds to a first serpentine ligand on the surface of a cell associated with the disorder; administering a second antibody coupled to a second factor, wherein the antibody binds to a second ligand on the surface of a cell, different from the first ligand; and wherein the first and second factors react to form a cytotoxic drug, thereby inhibiting the growth of the cell.

In a further embodiment, the invention provides a method for producing a specific cell lineage or organ type or an organism. For example, a totipotential germ cell

obtained by a method of the invention can be treated under conditions and for a time sufficient to produce the lineage, organ or organism. One non-limiting example of a use for a totipotential germ cell of the invention is for use in nuclear transplantation to produce new organisms.

In another embodiment, the invention provides a method of detecting a cell proliferative disorder in a sample including contacting the sample with a first antibody coupled to a factor wherein the antibody is reactive with an serpentine ligand on the surface of a cell having a cell proliferative disorder; contacting the sample with a second antibody coupled to a second factor, wherein the antibody is reactive to a second ligand on the surface of a cell; and wherein the antibody is reactive to a second ligand on the surface of a cell; and wherein the first and second factors react to form a reporter agent.

In another embodiment, the invention provides a method for detecting at least one variation in at least one serpentine polynucleotide in a cell, wherein the variation is indicative of a particular lineage. The method includes a) contacting a sample containing nucleic acid isolated from a first cell with at least one probe that hybrides to the at least one serpentine polynucleotide in the sample, wherein the serpentine polynucleotide is associated with a particular lineage or cell type, thereby forming a hybridization complex; and b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of a serpentine polynucleotide having a nucleic acid sequence complementary to the probe, and wherein the presence of the complex is indicative of a particular lineage or cell type. For example, one of skill in the art may select a particular protocadherin polynucleotide as a probe, wherein the protocadherin is associated with a particular cell type of lineage. The presence of the protocadherin, in combination with other cell surface markers, for example, is indicative of a particular lineage. Such a method allows

one of skill in the art to detect the switching of DNA when a cell moves from one lineage decision to another. It should be understood that one or more probes for several cell surface molecules are "multiplexed" in the method for identifying or detecting cell lineages or cell types. DNA or RNA analysis can be utilized in the method of the invention to detect DNA switching, for example by performing Southern Blot or Northern Blot analyses, respectively.

In yet another embodiment, the invention provides a method for producing a lineage-specific cell type comprising contacting a cell of a first lineage with a lineage determining effective amount of an agent or ligand for a time and under conditions such that the cell is committed to a second lineage, thereby producing a lineage-specific cell type. Examples of agents that can be used include an antibody, a ligand, a hormone, a growth factor, an antisense oligonucleotide or any combination thereof. Examples of growth factors or cytokines that may be included in the method of the invention include VEGF, FGF (e.g., FGF-2), CNTF, BDNF, IGF, interferons and interleukins (e.g., IL-4, IL-8).

In a further embodiment, the invention provides a method for maintaining a lineage-specific cell type comprising contacting the cell with a lineage maintaining effective amount of an agent or ligand such that the cell is committed to the lineage. Examples of regulatory proteins to which the agent, or ligand binds in producing or maintaining cell lineages, include Notch or Delta or related family members.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1. Hypothetical mechanism for the assembly of the precise topological map of glomeruli: A gradient of molecular affinities of olfactory receptors. About one thousand molecularly distinct glomeruli are arranged in a topologically precise map in the olfactory bulb. This map is bilaterally symmetrical but only one side is illustrated here. There are four distinct zones of glomeruli in the bulb (47-50), illustrated here in various shades of red, yellow, green, and blue. Gradients of colors on glomeruli within each zone are used to suggest an orderly gradient of molecular affinities of the individual receptors. A stream of migrating neurons originates in a specific fate-mapped region of the subventricular zone (22). Cells migrate as streams with the growth cones of each contacting the cell ahead (21). Colors and gradients are used again to suggest that receptors on each cell differ in an orderly way so that neighboring cells have receptors that bind with the highest affinity to each other. After reaching the olfactory bulb, cells change their direction of migration and move toward the surface of the bulb where they generate periglomerular cells (22). The dendrites of these cells then form the targets for incoming growth cones of olfactory nerve axons. Hundreds of olfactory neurons bearing the same, specific, olfactory receptor converge on a single pair of bilaterally symmetrical glomeruli (10-12). Their growth cones synapse with the dendrites of the periglomerular cells presumed to express the identical receptor. These homophilic interactions occur with the highest affinity. According to this hypothesis, receptors on neighboring glomeruli have closely related but different structures hence are bound with a slightly lower affinity. Mitral/tufted cells also synapse with glomeruli but are not illustrated here.

Figure 2. Diagram of a region of human chromosome 17 that codes for two olfactory receptors. This figure, based on the work of Glusman et al. (46), illustrates one of many sequenced regions of chromosomes that code for olfactory receptors and also contain numerous mobile elements. Note the pattern of elements near the upstream control elements of the two olfactory receptor coding regions (0R228 and OR 40) See the original publication for more details of this work. We hypothesize that some of these elements are used as genetic switches for the control of the expression of the thousand or more olfactory receptors.

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides a model for explaining the remarkable specificity of cell migration and tissue assembly that occurs throughout embryogenesis. The Area Code Hypothesis proposes that cells assemble organisms, including their brains and nervous systems, with the aid of a molecular addressing code that functions much like the country, area, regional and local portions of the telephone dialing system. The complexity of the information required to code cells for the construction of entire organisms is so enormous that we assume that the code must make combinatorial use of members of large multigene families. Such a system would reuse the same receptors as molecular digits in various regions of the embryo, thus greatly reducing the total number of genes required. Members of the very large families of serpentine receptors and vomeronasal (VNO) receptors fulfill the criteria proposed for area code molecules and serve as the last digits in such a code. Receptors of these families are expressed in many parts of developing embryos and suggest that they play a key functional role in cell recognition and targeting not only in the olfactory system but also throughout the brain and numerous other organs as they are assembled.

The present invention describes a number of uses that may be accomplished, utilizing the proposed hypothesis that serpentine receptors, and other seven-helical trans-membrane receptors (hereafter collectively "serpentine receptors"), code for cellular addresses in complex organisms. These receptors, which provide address and lineage markers for a broad range of cells and organs, can be used diagnostically and/or therapeutically, as well as to isolate and proliferate lineage-specific progenitor cells for organ regeneration or whole organism generation.

Receptors are classified into families and superfamilies on the basis of conserved structural features. It is generally believed that under selective pressure for organisms to acquire new biological functions, new receptor family members arose from duplication of existing receptor genes leading to the existence of multi-gene families. Family members thus contain vestiges of the ancestral gene, and these characteristic features can be exploited in the isolation and identification of additional family members. Some of the most well-known receptor superfamilies are the cytokine receptor superfamily, the protocadherin receptor family, the cell adhesion molecule (CAM) family, the seven transmembrane domain (7-TMD) receptor superfamily, and the steroid receptor superfamily. The cell surface display molecules referred to in the present invention are preferably "serpentine' receptors.

Based on the observation that numerous serpentine receptors are transcribed in a wide variety of organs and tissues, the present invention provides that such receptors have functions or utility outside of those specifically related to olfaction.

Further analysis of these surface serpentine receptor molecules reveals a pattern that parallels the cellular differentiation processes that occurs during embryogenesis, organogenesis, and limb formation. In addition, study of the DNA and RNA coding for these molecules, either in individual cells or within isolated tissues suggests characteristic, lineage-

specific genetic switching and rearrangements as lineage decisions are made by a cell. The switching that occurs at the DNA level is much like the switching that occurs in immunoglobulin DNA switching.

The characterization of surface components of cells, on a tissue by tissue basis, would be a daunting task. The present invention, however, provides a rapid and unifying mechanism to characterize tissues and even individual cells, according to the genetic organization and the display, or the lack of display, of serpentine receptors alone or in combination with other cell surface molecules.

DIAGNOSTIC USES

This is the first suggestion that serpentine receptors or the nucleic acids that code for serpentine receptors may be used to characterize or identify specific cell types. Binding agents such as ligands or antibodies, specific for serpentine receptors, are used for such identification and characterization. The preparation of polyclonal antibodies is well-known to those skilled in the art. See, for example, Green *et al.*, *Production of Polyclonal Antisera*, in IMMUNOCHEMICAL PROTOCOLS (Manson, ed.), pages 1-5 (Humana Press 1992); Coligan *et al.*, *Production of Polyclonal Antisera in Rabbits, Rats, Mice and Hamsters*, in CURRENT PROTOCOLS IN IMMUNOLOGY, section 2.4.1 (1992), which are hereby incorporated by reference.

One embodiment of the invention provides a method of obtaining a specific cell type or lineage. The method includes obtaining a sample of cells, contacting the cells with an agent, such as an antibody or a ligand specific for a cell surface marker such that the antibody or ligand binds to a cell in the sample and separating the cell that is bound by the antibody or ligand from the sample thereby obtaining a population of a specific cell type or lineage. The

cell population may be further purified by selecting for cells by expression of at least one additional marker associated with a specific cell type. For example, the additional marker may include CD-34, Thy-1, rho, Cdw109, protocadherins, and cell adhesion molecules, such as O-CAM, alone or in combination with other cell surface receptors. The method of the invention includes identifying a cell type by detecting expression of at least one serpentine receptor, wherein the presence or absence of the serpentine receptor is indicative of a cell type or lineage. In addition to analyzing the cell surface receptor polypeptides, one can also analyze the genetic fingerprint of the cell, e.g., identify changes in DNA as a result of switching or detect the presence or absence of RNA transcripts.

The preparation of monoclonal antibodies likewise is conventional. See, for example, Kohler & Milstein, *Nature* 256:495 (1975); Coligan *et al.*, sections 2.5.1-2.6.7; and Harlow *et al.*, ANTIBODIES: A LABORATORY MANUAL, page 726 (Cold Spring Harbor Pub. 1988), which are hereby incorporated by reference. Briefly, monoclonal antibodies can be obtained by injecting mice with a composition comprising an antigen/ligand, verifying the presence of antibody production by analyzing a serum sample, removing the spleen to obtain B lymphocytes, fusing the B lymphocytes with myeloma cells to produce hybridomas, cloning the hybridomas, selecting positive clones that produce antibodies to the antigen, and isolating the antibodies from the hybridoma cultures. Monoclonal antibodies can be isolated and purified from hybridoma cultures by a variety of well-established techniques. Such isolation techniques include affinity chromatography with Protein-A Sepharose, size-exclusion chromatography, and ion-exchange chromatography. *See, e.g.*, Coligan *et al.*, sections 2.7.1-2.7.12 and sections 2.9.1-2.9.3; Barnes *et al.*, *Purification of Immunoglobulin G (IgG)*, in METHODS IN MOLECULAR BIOLOGY, VOL. 10, pages 79-104 (Humana Press 1992).

Therapeutic applications for antibodies disclosed herein are also part of the present invention. Antibodies of the present invention may also be derived from subhuman primate antibody. General techniques for raising therapeutically useful antibodies in baboons can be found, for example, in Goldenberg *et al.*, International Patent Publication WO 91/11465 (1991) and Losman *et al.*, *Int. J. Cancer* 46:310 (1990), which are hereby incorporated by reference.

Alternatively, a therapeutically useful anti-olfactory antibody may be derived from a "humanized" monoclonal antibody. Humanized monoclonal antibodies are produced by transferring mouse complementarity determining regions from heavy and light variable chains of the mouse immunoglobulin into a human variable domain, and then substituting human residues in the framework regions of the murine counterparts. The use of antibody components derived from humanized monoclonal antibodies obviates potential problems associated with the immunogenicity of murine constant regions. General techniques for cloning murine immunoglobulin variable domains are described, for example, by Orlandi *et al.*, *Proc. Nat'l Acad. Sci. USA* 86:3833 (1989), which is hereby incorporated in its entirety by reference. Techniques for producing humanized monoclonal antibodies are described, for example, by Jones *et al.*, *Nature* 321: 522 (1986); Riechmann *et al.*, *Nature* 332: 323 (1988); Verhoeyen *et al.*, *Science* 239: 1534 (1988); Carter *et al.*, *Proc. Nat'l Acad. Sci. USA* 89: 4285 (1992); Sandhu, *Crit. Rev. Biotech.* 12: 437 (1992); and Singer *et al.*, *J. Immunol.* 150: 2844 (1993), which are hereby incorporated by reference.

Antibodies of the invention also may be derived from human antibody fragments isolated from a combinatorial immunoglobulin library. See, for example, Barbas *et al.*, METHODS: A COMPANION TO METHODS IN ENZYMOLOGY, VOL. 2, page 119 (1991); Winter *et al.*, *Ann. Rev. Immunol.* 12: 433 (1994), which are hereby incorporated by

reference. Cloning and expression vectors that are useful for producing a human immunoglobulin phage library can be obtained, for example, from STRATAGENE Cloning Systems (La Jolla, CA).

In addition, antibodies of the present invention may be derived from a human monoclonal antibody. Such antibodies are obtained from transgenic mice that have been "engineered" to produce specific human antibodies in response to antigenic challenge. In this technique, elements of the human heavy and light chain loci are introduced into strains of mice derived from embryonic stem cell lines that contain targeted disruptions of the endogenous heavy and light chain loci. The transgenic mice can synthesize human antibodies specific for human antigens and can be used to produce human antibody-secreting hybridomas. Methods for obtaining human antibodies from transgenic mice are described by Green *et al.*, *Nature Genet.* 7:13 (1994); Lonberg *et al.*, *Nature* 368:856 (1994); and Taylor *et al.*, *Int. Immunol.* 6:579 (1994), which are hereby incorporated by reference.

Antibody fragments of the present invention can be prepared by proteolytic hydrolysis of the antibody or by expression in *E. coli* of nucleic acid encoding the fragment. Antibody fragments can be obtained by pepsin or papain digestion of whole antibodies by conventional methods. For example, antibody fragments can be produced by enzymatic cleavage of antibodies with pepsin to provide a 5S fragment denoted $F(ab')_2$. This fragment can be further cleaved using a thiol reducing agent, and optionally a blocking group for the sulfhydryl groups resulting from cleavage of disulfide linkages, to produce 3.5S Fab' monovalent fragments. Alternatively, an enzymatic cleavage using pepsin produces two monovalent Fab' fragments and an Fc fragment directly. These methods are described, for example, by Goldenberg, U.S. patents No. 4,036,945 and No. 4,331,647, and references contained therein. These patents are hereby incorporated in their entireties by reference. *See*

also Nisonhoff *et al.*, *Arch. Biochem. Biophys.* 89:230 (1960); Porter, *Biochem. J.* 73:119 (1959); Edelman *et al.*, METHODS IN ENZYMOLOGY, VOL. 1, page 422 (Academic Press 1967); and Coligan *et al.* at sections 2.8.1-2.8.10 and 2.10.1-2.10.4.

Other methods of cleaving antibodies, such as separation of heavy chains to form monovalent light-heavy chain fragments, further cleavage of fragments, or other enzymatic, chemical, or genetic techniques may also be used, so long as the fragments bind to the antigen that is recognized by the intact antibody.

For example, Fv fragments comprise an association of V_H and V_L chains. This association may be noncovalent, as described in Inbar *et al.*, *Proc. Nat'l Acad. Sci. USA* 69:2659 (1972). Alternatively, the variable chains can be linked by an intermolecular disulfide bond or cross-linked by chemicals such as glutaraldehyde. *See, e.g.*, Sandhu, *supra.* Preferably, the Fv fragments comprise V_H and V_L chains connected by a peptide linker. These single-chain antigen binding proteins (sFv) are prepared by constructing a structural gene comprising nucleic acid sequences encoding the V_H and V_L domains connected by an oligonucleotide. The structural gene is inserted into an expression vector, which is subsequently introduced into a host cell such as *E. coli*. The recombinant host cells synthesize a single polypeptide chain with a linker peptide bridging the two V domains. Methods for producing sFvs are described, for example, by Whitlow *et al.*, METHODS: A COMPANION TO METHODS IN ENZYMOLOGY, VOL. 2, page 97 (1991); Bird *et al.*, *Science* 242:423-426 (1988); Ladner *et al.*, U.S. patent No. 4,946,778; Pack *et al.*, *Bio/Technology* 11: 1271-77 (1993); and Sandhu, *supra.*

Another form of an antibody fragment is a peptide coding for a single complementarity-determining region (CDR). CDR peptides ("minimal recognition units") can be obtained by constructing genes encoding the CDR of an antibody of interest. Such

genes are prepared, for example, by using the polymerase chain reaction to synthesize the variable region from RNA of antibody-producing cells. See, for example, Larrick *et al.*, METHODS: A COMPANION TO METHODS IN ENZYMOLOGY, VOL. 2, page 106 (1991).

The term "antibody" as used in this invention includes intact molecules as well as fragments thereof, such as Fab, F(ab')2, and Fv which are capable of binding to an epitopic determinant present in Bin1 polypeptide. Such antibody fragments retain some ability to selectively bind with its antigen or receptor.

As used in this invention, the term "epitope" refers to an antigenic determinant on an antigen to which the paratope of an antibody binds. Epitopic determinants usually consist of chemically active surface groupings of molecules such as amino acids or sugar side chains and usually have specific three dimensional structural characteristics, as well as specific charge characteristics.

Antibodies which bind to serpentine receptor polypeptides can be prepared using an intact polypeptide or fragments containing small peptides of interest as the immunizing antigen. For example, it may be desirable to produce antibodies that specifically bind to the extracellular loop, or the N- or C-terminal or other domains of a serpentine receptor. The polypeptide or peptide used to immunize an animal which is derived from translated cDNA or chemically synthesized which can be conjugated to a carrier protein, if desired. Such commonly used carriers which are chemically coupled to the immunizing peptide include keyhole limpet hemocyanin (KLH), thyroglobulin, bovine serum albumin (BSA), and tetanus toxoid.

The polyclonal or monoclonal antibodies can be further purified, for example, by binding to and elution from a matrix to which the polypeptide or a peptide to which the

antibodies were raised is bound. Those of skill in the art will know of various techniques common in the immunology arts for purification and/or concentration of polyclonal antibodies, as well as monoclonal antibodies (See for example, Coligan, *et al.*, Unit 9, Current Protocols in Immunology, Wiley Interscience, 1994, incorporated by reference). One can also couple or link the antibody to additional factors such as other peptides, polypeptides or chemical structures.

It is also possible to use the anti-idiotype technology to produce invention monoclonal antibodies which mimic an epitope. For example, an anti-idiotypic monoclonal antibody made to a first monoclonal antibody will have a binding domain in the hypervariable region which is the "image" of the epitope bound by the first monoclonal antibody.

In addition, genetic probes, such as DNA or RNA polynucleotides, can also be made to identify the extent of genetic rearrangement in serpentine receptor genes, and thereby characterize or identify the cells in question.

Detection of a nucleic acid encoding an serpentine receptor may be performed by standard methods such as size fractionating the nucleic acid. Methods of size fractionating the DNA and RNA are well known to those of skill in the art, such as by gel electrophoresis, including polyacrylamide gel electrophoresis (PAGE). For example, the gel may be a denaturing 7 M or 8 M urea-polyacrylamide-formamide gel. Size fractionating the nucleic acid may also be accomplished by chromatographic methods known to those of skill in the art.

The detection of polynucleotides may optionally be performed by using radioactively labeled probes. Any radioactive label may be employed which provides an adequate signal. Other labels include ligands, which can serve as a specific binding pair member for a labeled ligand, and the like. The labeled preparations are used to probe nucleic acid by the Southern

or Northern hybridization techniques, for example. Nucleotides obtained from samples are transferred to filters that bind polynucleotides. After exposure to the labeled nucleic acid probe, which will hybridize to nucleotide fragments containing target nucleic acid sequences, the binding of the radioactive probe to target nucleic acid fragments is identified by autoradiography (see *Genetic Engineering, 1,* ed. Robert Williamson, Academic Press (1981), pp. 72-81). The particular hybridization technique is not essential to the invention. Hybridization techniques are well known or easily ascertained by one of ordinary skill in the art. As improvements are made in hybridization techniques, they can readily be applied in the method of the invention.

The polynucleotides encoding the desired polypeptide may be amplified before detecting. The term "amplified" refers to the process of making multiple copies of the nucleic acid from a single polynucleotide molecule. The amplification of polynucleotides can be carried out *in vitro* by biochemical processes known to those of skill in the art. The amplification agent may be any compound or system that will function to accomplish the synthesis of primer extension products, including enzymes. Suitable enzymes for this purpose include, for example, *E. coli* DNA polymerase I, Taq polymerase, Klenow fragment of *E. coli* DNA polymerase I, T4 DNA polymerase, other available DNA polymerases, polymerase muteins, reverse transcriptase, ligase, and other enzymes, including heat-stable enzymes (*i.e.,* those enzymes that perform primer extension after being subjected to temperatures sufficiently elevated to cause denaturation). Suitable enzymes will facilitate combination of the nucleotides in the proper manner to form the primer extension products that are complementary to each mutant nucleotide strand. Generally, the synthesis will be initiated at the 3' end of each primer and proceed in the 5' direction along the template strand, until synthesis terminates, producing molecules of different lengths. There may be

amplification agents, however, that initiate synthesis at the 5' end and proceed in the other direction, using the same process as described above. In any event, the method of the invention is not to be limited to the embodiments of amplification described herein.

One method of *in vitro* amplification which can be used according to this invention is the polymerase chain reaction (PCR) described in U.S. Patent Nos. 4,683,202 and 4,683,195. The term "polymerase chain reaction" refers to a method for amplifying a DNA base sequence using a heat-stable DNA polymerase and two oligonucleotide primers, one complementary to the (+)-strand at one end of the sequence to be amplified and the other complementary to the (-)-strand at the other end. Because the newly synthesized DNA strands can subsequently serve as additional templates for the same primer sequences, successive rounds of primer annealing, strand elongation, and dissociation produce rapid and highly specific amplification of the desired sequence. The polymerase chain reaction is used to detect the presence of polynucleotides encoding cell surface receptors in the sample. Many polymerase chain methods are known to those of skill in the art and may be used in the method of the invention. For example, DNA can be subjected to 30 to 35 cycles of amplification in a thermocycler as follows: 95°C for 30 sec, 52° to 60°C for 1 min, and 72°C for 1 min, with a final extension step of 72°C for 5 min. For another example, DNA can be subjected to 35 polymerase chain reaction cycles in a thermocycler at a denaturing temperature of 95ºC for 30 sec, followed by varying annealing temperatures ranging from 54-58ºC for 1 min, an extension step at 70ºC for 1 min and a final extension step at 70ºC.

The primers for use in amplifying the polynucleotides of the invention may be prepared using any suitable method, such as conventional phosphotriester and phosphodiester methods or automated embodiments thereof so long as the primers are capable of hybridizing to the polynucleotides of interest. One method for synthesizing oligonucleotides on a

modified solid support is described in U.S. Patent No. 4,458,066. The exact length of primer will depend on many factors, including temperature, buffer, and nucleotide composition. The primer must prime the synthesis of extension products in the presence of the inducing agent for amplification.

Primers used according to the method of the invention are complementary to each strand of nucleotide sequence to be amplified. The term "complementary" means that the primers must hybridize with their respective strands under conditions which allow the agent for polymerization to function. In other words, the primers that are complementary to the flanking sequences hybridize with the flanking sequences and permit amplification of the nucleotide sequence. Preferably, the 3' terminus of the primer that is extended has perfectly base paired complementarity with the complementary flanking strand.

Those of ordinary skill in the art will know of various amplification methodologies which can also be utilized to increase the copy number of target nucleic acid. The polynucleotides detected in the method of the invention can be further evaluated, detected, cloned, sequenced, and the like, either in solution or after binding to a solid support, by any method usually applied to the detection of a specific nucleic acid sequence such as another polymerase chain reaction, oligomer restriction (Saiki *et al., Bio/Technology 3*: 1008-1012 (1985)), allele-specific oligonucleotide (ASO) probe analysis (Conner *et al., Proc. Natl. Acad. Sci. USA 80*: 278 (1983), oligonucleotide ligation assays (OLAs) (Landegren *et al., Science 241*: 1077 (1988)), RNAse Protection Assay and the like. Molecular techniques for DNA analysis have been reviewed (Landegren *et al, Science, 242:* 229-237 (1988)). Following DNA amplification, the reaction product may be detected by Southern blot analysis, without using radioactive probes. In such a process, for example, a small sample of DNA containing a the polynucleotides obtained from the tissue or subject are amplified, and

analyzed via a Southern blotting technique. The use of non-radioactive probes or labels is facilitated by the high level of the amplified signal. In a one embodiment of the invention, one nucleoside triphosphate is radioactively labeled, thereby allowing direct visualization of the amplification product by autoradiography. In another embodiment, amplification primers are fluorescent labeled and run through an electrophoresis system. Visualization of amplified products is by laser detection followed by computer assisted graphic display. Simple visualization of a gel containing the separated products may be utilized to determine the presence of a polynucleotide. However, other methods known to those skilled in the art may also be used, for example scanning densitometry, computer aided scanning and quantitation.

Polynucleotides encoding serpentine receptors may be identified by nucleic acid hybridization techniques. In nucleic acid hybridization reactions, the conditions used to achieve a particular level of stringency will vary, depending on the nature of the nucleic acids being hybridized. For example, the length, degree of complementarity, nucleotide sequence composition (*e.g.*, GC v. AT content), and nucleic acid type (*e.g.*, RNA v. DNA) of the hybridizing regions of the nucleic acids can be considered in selecting hybridization conditions. An additional consideration is whether one of the nucleic acids is immobilized, for example, on a filter.

An example of progressively higher stringency conditions is as follows: 2 x SSC/0.1% SDS at about room temperature (hybridization conditions); 0.2 x SSC/0.1% SDS at about room temperature (low stringency conditions); 0.2 x SSC/0.1% SDS at about 42°C (moderate stringency conditions); and 0.1 x SSC at about 68°C (high stringency conditions). Washing can be carried out using only one of these conditions, *e.g.*, high stringency conditions, or each of the conditions can be used, *e.g.*, for 10-15 minutes each, in the order listed above, repeating any or all of the steps listed. However, as mentioned above, optimal

conditions will vary, depending on the particular hybridization reaction involved, and can be determined empirically.

Biological chips or arrays are useful in a variety of screening techniques for obtaining information about serpentine receptor display on cell surfaces. Arrays of nucleic acid probes can be used to extract sequence information from, for example, nucleic acid samples. The samples are exposed to the probes under conditions that allow hybridization. The arrays are then scanned to determine to which probes the sample molecules have hybridized. One can obtain sequence information by careful probe selection and using algorithms to compare patterns of hybridization and non-hybridization. This method is useful for sequencing nucleic acids, as well as sequence checking. For example, the method is useful in diagnostic screening for genetic diseases or for the presence and/or identity of a particular pathogen or a strain of pathogen. For example, there are various strains of HIV, the virus that causes AIDS. Some of them have become resistant to current AIDS therapies. Diagnosticians can use DNA arrays to examine a nucleic acid sample from the virus to determine what strain it belongs to. In the same way, the genetic fingerprint including nucleic acid coding for serpentine receptors, can be compared with nucleic acid samples extracted from different cell samples.

The biological chip plates used in the methods of this invention include biological chips. The array of probe sequences can be fabricated on the biological chip according to the pioneering techniques disclosed in U.S. Pat. No. 5,143,854, PCT WO 92/10092, PCT WO 90/15070, or U.S. application Ser. Nos. 08/249,188, 07/624,120, and 08/082,937. The combination of photolithographic and fabrication techniques may, for example, enable each probe sequence ("feature") to occupy a very small area ("site" or "location") on the support. In some embodiments, this feature site may be as small as a few microns or even a single molecule. For example, a probe array of 0.25 mm.sup.2 (about the size that would fit in a

well of a typical 96-well microtiter plate) could have at least 10, 100, 1000, 10_4, 10_5 or 10_6 features. In an alternative embodiment, such synthesis is performed according to the mechanical techniques disclosed in U.S. Pat. No. 5,384,261, incorporated herein by reference. Sensitive analysis of serpentine receptor nucleic acid can also be performed as described by Clinical Microsystems, using AC to detect minute changes in electron flow in dsDNA after DNA fragments hybridize to an array of DNA on a chip.

Cells which contain the nucleic acid sequence encoding serpentine receptors and express serpentine receptors may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein.

The presence of polynucleotide sequences encoding serpentine receptors can be detected by DNA-DNA or DNA-RNA hybridization or amplification using probes or fragments or fragments of polynucleotides encoding serpentine receptors. Nucleic acid amplification based assays involve the use of oligonucleotides or oligomers based on the sequences encoding serpentine receptors to detect cells containing DNA or RNA encoding serpentine receptors.

A variety of protocols for detecting and measuring the expression of serpentine receptors, using either polyclonal or monoclonal antibodies specific for the protein are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on serpentine receptors can be used, but a competitive binding assay may

be employed. These and other assays are described, among other places, in Hampton, R. et al. (1990; Serological Methods, a Laboratory Manual, APS Press, St Paul, Minn.) and Maddox, D. E. et al. (1983; J. Exp. Med. 158:1211-1216).

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding serpentine receptors include oligolabeling, nick translation, end-labeling or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding serpentine receptors, or any fragments thereof may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits (Pharmacia & Upjohn, (Kalamazoo, Mich.); Promega (Madison Wis.); and U.S. Biochemical Corp., Cleveland, Ohio). Suitable reporter molecules or labels, which may be used for ease of detection, include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

In another embodiment, antibodies which specifically bind serpentine receptors may be used for the diagnosis of conditions or diseases characterized by expression of specific serpentine receptors, or in assays to monitor patients being treated with serpentine receptors, agonists, antagonists or inhibitors. The antibodies useful for diagnostic purposes may be prepared in the same manner as those described for therapeutics. Diagnostic assays for serpentine receptors include methods which utilize the antibody and a label to detect serpentine receptors in human body fluids or extracts of cells or tissues. The antibodies may

be used with or without modification, and may be labeled by joining them, either covalently or non-covalently, with a reporter molecule. A wide variety of reporter molecules which are known in the art may be used, several of which are described above.

A variety of protocols including ELISA, RIA, and FACS for measuring serpentine receptors are known in the art and provide a basis for diagnosing levels of serpentine receptors expression. Normal or standard values for serpentine receptors expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, preferably human, with antibody to serpentine receptors under conditions suitable for complex formation. The amount of standard complex formation may be quantified by various methods, but preferably by photometric, means. Quantities of serpentine receptors expressed in subject, control and disease, samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding serpentine receptors may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantitate gene expression in biopsied tissues in which expression of serpentine receptors may be correlated with disease. The diagnostic assay may be used to distinguish between absence, presence, and excess expression of serpentine receptors, and to monitor regulation of serpentine receptors levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding serpentine receptors or closely related molecules, may be used to identify nucleic acid sequences which encode

serpentine receptors. The specificity of the probe, whether it is made from a highly specific region, e.g., 10 unique nucleotides in the 5' regulatory region, or a less specific region, e.g., especially in the 3' coding region, and the stringency of the hybridization or amplification (maximal, high, intermediate, or low) will determine whether the probe identifies only naturally occurring sequences encoding serpentine receptors, alleles, or related sequences.

Probes may also be used for the detection of related sequences, and should preferably contain at least 50% of the nucleotides from any of the serpentine receptors encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and derived from the nucleotide sequence (e.g., mRNA) of serpentine receptors or from genomic sequence including promoter, enhancer elements, and introns of the naturally occurring serpentine receptors.

Means for producing specific hybridization probes for DNAs encoding serpentine receptors include the cloning of nucleic acid sequences encoding serpentine receptors or serpentine receptors derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, radionuclides such as 32P or 35S, or enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding serpentine receptors may involve the use of PCR. Such oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably consist of two nucleotide sequences, one with sense orientation (5' to 3') and another with antisense (3' to 5'), employed under optimized conditions for identification of a specific gene or

condition. The same two oligomers, nested sets of oligomers, or even a degenerate pool of oligomers may be employed under less stringent conditions for detection and/or quantitation of closely related DNA or RNA sequences.

Methods which may also be used to quantitate the expression of serpentine receptors include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and standard curves onto which the experimental results are interpolated (Melby, P. C. et al. (1993) J. Immunol. Methods, 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 229-236). The speed of quantitation of multiple samples may be accelerated by running the assay in an ELISA format where the oligomer of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, an oligonucleotide derived from any of the polynucleotide sequences described herein may be used as a target in a microarray. The microarray can be used to monitor the expression level of large numbers of genes simultaneously (to produce a transcript image), and to identify genetic variants, mutations and polymorphisms. This information will be useful in determining gene function, understanding the genetic basis of disease, diagnosing disease, and in developing and monitoring the activity of therapeutic agents (Heller, R. et al. (1997) Proc. Natl. Acad. Sci. 94:2150-55).

In one embodiment, the microarray is prepared and used according to the methods described in PCT application WO95/11995 (Chee et al.), Lockhart, D. J. et al. (1996; Nat. Biotech. 14: 1675-1680) and Schena, M. et al. (1996; Proc. Natl. Acad. Sci. 93: 10614-10619), all of which are incorporated herein in their entirety by reference.

The microarray is preferably composed of a large number of unique, single-stranded nucleic acid sequences, usually either synthetic antisense oligonucleotides or fragments of cDNAs, fixed to a solid support. The oligonucleotides are preferably about 6-60 nucleotides

in length, more preferably 15-30 nucleotides in length, and most preferably about 20-25 nucleotides in length. For a certain type of microarray, it may be preferable to use oligonucleotides which are only 7-10 nucleotides in length. The microarray may contain oligonucleotides which cover the known 5', or 3', sequence, sequential oligonucleotides which cover the full length sequence; or unique oligonucleotides selected from particular areas along the length of the sequence. Polynucleotides used in the microarray may be oligonucleotides that are specific to a gene or genes of interest in which at least a fragment of the sequence is known or that are specific to one or more unidentified cDNAs which are common to a particular cell type, developmental or disease state.

In order to produce oligonucleotides to a known sequence for a microarray, the gene of interest is examined using a computer algorithm which starts at the 5' or more preferably at the 3' end of the nucleotide sequence. The algorithm identifies oligomers of defined length that are unique to the gene, have a GC content within a range suitable for hybridization, and lack predicted secondary structure that may interfere with hybridization. In certain situations it may be appropriate to use pairs of oligonucleotides on a microarray. The "pairs" will be identical, except for one nucleotide which preferably is located in the center of the sequence. The second oligonucleotide in the pair (mismatched by one) serves as a control. The number of oligonucleotide pairs may range from two to one million. The oligomers are synthesized at designated areas on a substrate using a light-directed chemical process. The substrate may be paper, nylon or other type of membrane, filter, chip, glass slide or any other suitable solid support.

In another aspect, an oligonucleotide may be synthesized on the surface of the substrate by using a chemical coupling procedure and an ink jet application apparatus, as described in PCT application WO95/25 1116 (Baldeschweiler et al.) which is incorporated

herein in its entirety by reference. In another aspect, a "gridded" array analogous to a dot (or slot) blot may be used to arrange and link cDNA fragments or oligonucleotides to the surface of a substrate using a vacuum system, thermal, UV, mechanical or chemical bonding procedures. An array, such as those described above, may be produced by hand or by using available devices (slot blot or dot blot apparatus), materials (any suitable solid support), and machines (including robotic instruments), and may contain 8, 24, 96, 384, 1536 or 6144 oligonucleotides, or any other number between two and one million which lends itself to the efficient use of commercially available instrumentation. In a similar manner, ligands or antibodies, or their derivatives, can be utilized to generate arrays.

In order to conduct sample analysis using a microarray, the RNA or DNA from a biological sample is made into hybridization probes. The mRNA is isolated, and cDNA is produced and used as a template to make antisense RNA (aRNA). The aRNA is amplified in the presence of fluorescent nucleotides, and labeled probes are incubated with the microarray so that the probe sequences hybridize to complementary oligonucleotides of the microarray. Incubation conditions are adjusted so that hybridization occurs with precise complementary matches or with various degrees of less complementarity. After removal of nonhybridized probes, a scanner is used to determine the levels and patterns of fluorescence. The scanned images are examined to determine degree of complementarity and the relative abundance of each oligonucleotide sequence on the microarray. Analysis of microarrays of ligands and or antibodies can be performed in a similar manner. The biological samples may be obtained from any bodily fluids (such as blood, urine, saliva, phlegm, gastric juices, etc.), cultured cells, biopsies, or other tissue preparations. A detection system may be used to measure the absence, presence, and amount of hybridization or binding for all of the distinct molecules simultaneously. This data may be used for large scale correlation studies on the sequences,

mutations, variants, or polymorphisms among samples.

In another embodiment of the invention, the nucleic acid sequences which encode serpentine receptors may also be used to generate hybridization probes which are useful for mapping the naturally occurring genomic sequence and for detecting differences in the sequence that might be indicative of a lineage. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome or to artificial chromosome constructions, such as human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions or single chromosome cDNA libraries as reviewed in Price, C. M. (1993) Blood Rev. 7:127-134, and Trask, B. J. (1991) Trends Genet. 7:149-154.

Fluorescent *in situ* hybridization (FISH as described in Verma et al. (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York, N.Y.) may be correlated with other physical chromosome mapping techniques and genetic map data. Examples of genetic map data can be found in various scientific journals or at Online Mendelian Inheritance in Man (OMIM). Correlation between the location of the gene encoding serpentine receptors on a physical chromosomal map and a specific disease, or predisposition to a specific disease, may help delimit the region of DNA associated with a particular cell lineage. The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between cell lineages for diagnostic, therapeutic or other applications as discusssed throughout the specification.

In situ hybridization of chromosomal preparations and physical mapping techniques such as linkage analysis using established chromosomal markers may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian

species, such as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms, or parts thereof, by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, for example, AT to 11q22-23 (Gatti, R. A. et al. (1988) Nature 336:577-580), any sequences mapping to that area may represent associated or regulatory genes for further investigation. The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc. among normal, carrier, or affected individuals.

THERAPEUTIC USES

The "area code" hypothesis holds that localized tissues, even at the level of individual cells, display a unique receptor pattern. This fact could allow for exquisitely specific therapeutic strategies for the treatment of cancer, for the delivery of specific genetic therapies, for specific immunoregulation or immunosuppression and others.

The following is an example of how this might be used for cancer therapy. There are a number of molecules that, in isolation, are non-toxic. When combined with other non-toxic molecules, the combination is toxic. Imagine one such molecule, targeted by means of, say, an antibody to a specific serpentine receptor characteristic of a specific lineage (e.g., the particular B-cell lineage associated with a patient's lymphoma). The molecule would be drawn not only to the specific receptor site on those B-lymphoma cells (which is what you want) but also to other sites within the body (which you don't want). Then, if a second molecule (non-toxic unless combined with the first), likewise targeted to another surface

determinant of the lymphoma, is introduced, it finds the lymphoma cells and other, different cells. Only the cells that receive both molecules (lymphoma cells) are delivered a toxic dose, thereby reducing non-specific toxicity of the cancer drugs.

Such a scheme could also be used in genetic therapy approaches, with specific genetic sequences carrying enabling and coding functions delivered independently to different molecules of the serpentine receptor address, so that the genetic therapy is targeted appropriately. Also, complementary strands of RNA could be delivered independently in order to inhibit specific genes, since it is known that dsRNA can block gene transcription in ways that ssRNA (in antisense orientation) does not.

Another method for deliver of agents or compounds to cells can be accomplished by the use of multiplexed ligands by adjusting binding constants appropriately such that ligands bind only to cells of interest. Multiple binding or "multiplexing" allows one to target a specific cell type (e.g., a germ cell), having all of the "digits" of the "phone number". Furthermore, additional ligands, with or without a targeting sequence, can be utilized in assisting delivery of the agents to the nucleus of the cell. Agents or compounds include genes for gene therapy, drugs, toxins, small molecules, peptides, antibodies or fragments thereof and the like.

In addition to the serpentine receptors, it might be desirable to modulate regulatory proteins in the cell. Examples of such families of proteins include the Drosophila Notch protein, the Nematode lin-12 and glp-1 proteins, and the closely related vertebrate homologs, Motch (mouse Notch), Xotch (Xenopus Notch), rat Notch, and TAN 1 (human Notch), all of which are membrane bound receptor molecules that control the specification of cell fate for a variety of cell types early in embryogenesis (Rebay et al., 1991, Cell 67:687; Hutter and Schnabel, 1994, Development 120:2051; Del Amo et al 1992, Development 115:737;

Reaume et al. 1992 Develop. Biol. 154:377; and Ellisen et al., 1991, Cell 66:649). Specific EGF-like repeats in the Notch receptors are binding sites that attach to protein ligands leading to signal transduction (Rebay et al., 1991 Cell 67:687; Couso and Arias, 1994, Cell 79:259; Fortini and Artavanis-Tsakonas, 1994, Cell 79:273; Henderson et al., 1994, Development 120:2913). Extracellular matrix proteins such as thrombospondin, entactin, tenascin and laminin play key roles in morphogenesis by providing the physical scaffold to which cells attach to form and maintain tissue morphologies (Frazier, 1987, J. Cell. Biol. 105:625; Taraboletti et al., 1990, J. Cell. Biol. 111:765; Ekblom et al., 1994, Development 120:2003). It may also be desirable to modulate the Delta and/or Notch protein, which have been associated with cell fate. Notch in its active form, i.e., the form that mediates signal transduction and binds Notch ligands such as Delta (WO99/04746). Notch and Delta as well as related family members play a key role in switching events that mediate cell lineage decisions.

SCREENING FOR ANTAGONISTS OR AGONISTS OF TRANSMEMBRANE RECEPTORS

In another embodiment of the invention, serpentine receptors, their catalytic or immunogenic fragments or oligopeptides thereof, can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes, between serpentine receptors and the agent being tested, may be measured.

Another technique for drug screening which may be used provides for high throughput screening of compounds having suitable binding affinity to the protein of interest as described in published PCT application WO84/03564. In this method, as applied to

serpentine receptors large numbers of different small test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. The test compounds are reacted with serpentine receptors, or fragments thereof, and washed. Bound serpentine receptors is then detected by methods well known in the art. Purified serpentine receptors can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive screening assays in which ligands or other molecules capable of binding serpentine receptors specifically compete with a test compound or ligand for binding serpentine receptors. In this manner, the ligand or test compound can be used to detect the presence of any molecule which shares one or more antigenic or binding determinants with serpentine receptors.

In additional embodiments, the nucleotide sequences which encode serpentine receptors may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Regeneration of Tissue or Organs

Progenitor cells that are committed to being a specific cell type but still capable of further differentiation, including totipotential and pluripotential progenitor cells such as germ cells and mesenchymal stem cells, respectively, and more tissue specific progenitor cells such as chondrocytes, display specific serpentine or other receptors that have are characteristic of each lineage. The cell surface display of these codes can be used to identify cell-specific lineages. The importance of progenitor cells has been recognized already in some fields of

therapy, including tissue engineering, bone marrow abalation therapies etc. For example, progenitor cell lines isolated from bone marrow or circulating blood have been used to re-populate the hematopoietic system in individuals whose bone marrow is ablated and then reconstituted in bone marrow transplantation procedures. Certain neurological defects, such as Parkinson's disease and others, have been cured or ameliorated through the transplantation of fetal or immature tissues. These results have been made possible by a re-growth and re-differentiation of tissue originating from progenitor cells.

Utilizing cell surface address codes, including the serpentine receptors that characterize the surface of specific progenitor cells, these cells can be isolated by a number of cell selection techniques (FACS, immunomagnetic beads, others). Such selection techniques can include both positive selection, for example identifying and removing the cell of interest from a population, as well as negative selection, removal of the positive cells from the population leaving only the negative cells. Negative selection may prove useful in isolating cells that have yet to differentiate sufficiently to express a particular serpentine receptor. Further, an understanding of both the surface characteristics and also the genetic switching processes relating to serpentine receptors will be useful in the development of cell culture techniques to maintain and propagate such cells in their progenitor state. Purified progenitor cells are likely to become important therapeutic moieties in the treatment of disease and deficiencies.

Serpentine receptors and other transmembrane receptors are believed to function as the last digits in a cell surface code.

Data that we have obtained by searching the genome databases have provided us with evidence suggesting that the very large families of seven transmembrane receptors, including the serpentine receptors may indeed be used in a combinatorial array with other cell surface

address molecules during the assembly of many tissues in addition to the olfactory regions. Such molecules therefore have many of the properties expected for area code molecules.

The invention will now be described in greater detail by reference to the following non-limiting examples.

EXAMPLES

Internet Grateful Med and SciSearch (ISI) databases were used for retrieval of bibliographic information. Large numbers of references including abstracts were downloaded into Procite 4 (ISI) for further searching and analysis locally as well as for formatting references. The online resources available through The National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov) were used extensively in this work. The information that is reported in Table 1 was obtained by searching the dbEST database using the text string: olfactory AND receptor and the names of the tissues indicated in Table 1. The retrieved and curated information included in Table 1 represents only a partial list. The quality of the sequence data varied widely as is normal for the expressed sequence tags. Nevertheless, it was clear that this approach provided a great deal of useful information on the expression of serpentine receptor genes in a large number of different tissues. Only the retrieved sequences that are related to known serpentine receptors are included in Table 1. Other informative searches used known amino acid sequences of specific serpentine receptors from various species to retrieve expressed sequence tags. For these studies, BLAST 2.0 (Gapped BLAST and Graphical Viewer) with the advanced BLAST option was used. The tblastn program was used to search the dbEST database.

Typically, nucleic acid sequence information for a desired receptor or other protein can be located in one of several public databases, e.g., Genbank, EMBL, SwissProt, and PIR,

or in biological related journal publications. Thus, one of skill in the art would have access to nucleic acid sequence information for virtually all known genes. Those of skill in the art can either obtain the corresponding nucleic acid molecule directly from a public depository or the institution that published the sequence. Alternatively, once the nucleic acid sequence encoding a desired protein has been ascertained, the skilled artisan can employ routine methods, e.g., polymerase chain reaction (PCR) amplification to isolate the desired nucleic acid molecule from the appropriate nucleic acid library. Thus all known nucleic acids encoding proteins of interest, e.g., serpentine receptors, are available for use in the methods and products described herein.

EXAMPLE 1

OLFACTORY NEURONS EACH EXPRESS A SINGLE RECEPTOR AND USE THAT RECEPTOR TO TARGET A SPECIFIC PAIR OF BILATERALLY SYMMETRICAL GLOMERULI

Recent research including the elegant experiments by Mombaerts et al. (9,10) has shown that the olfactory receptors themselves do in fact play an important role in axonal targeting as their processes extend from the olfactory epithelium to specific glomeruli in the olfactory bulb. Neurons that express the same receptor gene but are dispersed in the olfactory epithelium target their processes to a single pair of bilaterally symmetrical glomeruli (11,12; see Fig. 1). There are one thousand or so different genes that code for olfactory receptors. About the same number of glomeruli are arranged in a precise, topologically ordered array in each of the two sides of the olfactory bulb. These serve as highly specific targets for the growth cones of the olfactory neurons, each expressing a single receptor gene. Because these olfactory receptors bear the hallmarks of the proposed area code molecules, it seemed

appropriate to ask if they might be expressed in other parts of the developing embryo (and adult) as expected for such molecular codes.

A search of the genome and literature databases revealed a remarkable number of examples of these genes expressed in tissues other than the olfactory system. Axons expressing VNO receptors are believed to target the accessory olfactory bulb with similar high precision and they too are assumed to play a role in cell targeting.

EXAMPLE 2

EXPRESSION OF MEMBERS OF THESE FAMILIES OF RECEPTORS IN TISSUES OTHER THAN THE OLFACTORY EPITHELIUM

Expressed sequence tags are being entered into the dbEST database at a rapid rate and now represent an important new resource for the study of gene expression. The cDNA samples used for these sequencing studies are obtained from a wide variety of tissues, developmental stages and organisms. The data vary in quality but nevertheless provide a rich source of information. A search of dbEST revealed many examples of the expression of olfactory receptor genes expressed in tissues other than the olfactory system. A partial set of results from this study is summarized in Table 1. Surprisingly, these genes are expressed in liver, lung, colon, testis, ovary, uterus, prostate, thyroid, brain and many other tissues and tumors. In addition, a search of the bibliographic databases revealed several publications dealing with the expression of olfactory receptors in a few tissues (13-15).

The original area code paper reviews a number of systems in which cell migration plays a role in organogenesis. The embryonic heart is a particularly interesting example of an organ that is assembled using migrating cells that coalesce and construct the tissue with great precision. In pursuing the notion that serpentine receptors can act as receptors in an area code

system we were therefore gratified to find in our searches of dbEST that specific olfactory receptors are indeed expressed in the embryonic heart. A publication was also found that provides further evidence for such expression (13). One olfactory receptor, OL1, was studied in detail and the data, including in situ hybridization studies, seem very convincing. The authors further state that other olfactory receptors are also expressed in the embryonic heart but give no data. It will be most interesting to learn the extent, timing, and topography of the expression of these receptors in the embryonic heart and also in the many other organs where they are expressed.

The widespread expression of members of the serpentine receptor family in numerous organ systems obviously supports the hypothesis that the receptors perform functions other than the recognition of olfactants. Since these receptors play a dual role as receptors for molecules in the olfactory epithelium and as cell surface addressing molecules that aid in the assembly of the olfactory bulb, one obvious notion is that they may also play a dual role in other parts of the embryo. The possibility of the combined functions of cell-cell recognition and organ construction, and also as cell surface receptors for many classes of small molecules, represents an extremely provocative concept when considering the roles of these very large families of genes. Another surprising consequence of this notion is that some of the very widely expressed receptors of the calcium sensing and metabotrophic glutamate families (found in the VNO/accessory olfactory system) may also have dual functions and thus play a role in cellular addressing during development. One would certainly not anticipate or postulate a dual role for these receptor classes if members of these families were not functional in the VNO olfactory system as receptors for pheromones and other small molecules and for targeting the accessory olfactory bulb (16-20).

Assembly of the Olfactory bulb: A Model for other Parts of the Brain and Embryo. As discussed above, several research groups agree that olfactory neurons expressing the same serpentine receptor, from among the one thousand or so total receptors, converge on a single pair of glomeruli in the olfactory bulb. A logical consequence of this fact is that each glomerulus in one of the bilaterally symmetrical olfactory lobes has a unique address on the fixed topological map of the olfactory bulb. There are about one thousand distinct addresses in each lobe. Furthermore, the maps are the same in each of the inbred individuals and they are believed to be "hardwired" by genetic programs that control development. It has been determined that the targets are established during embryogenesis. When the growth cones of olfactory neurons start entering the olfactory bulb, the targets await. It follows that the assembly of this target structure must itself use a very sophisticated molecular addressing system during embryogenesis and then display molecules that provide the topologically precise, distinct targets for olfactory nerve growth cones.

The subventricular zone, a considerable distance posterior to the region where the olfactory bulb is formed, is the birthplace of neuronal precursor cells that are destined to form the olfactory bulb. Topological fate maps of this region reveal various specific positions of cells that are destined to generate distinct parts of the forebrain. A small region in the extreme anterior of the subventricular zone is the source of cells that will begin the migration to the region where the olfactory bulb is assembled (21,22; see Fig. 1). We assume that migratory cells are generated in an ordered fashion from these precursor cells and that the order of birth of daughter cells relates to their ultimate position in the topology of the olfactory bulb. As such cells are born they begin migrating along a narrow tube-like pathway bounded by glial cells but, unlike other regions of the embryonic brain, no radial glial processes are seen. The migrating spindle shaped cells remain in contact with neighboring cells in front, beside and

behind and migrate as a stream only a few cells in diameter (21). Cell division continues while they migrate and maintain contacts. As cells in this stream reach the inner region of the developing olfactory bulb some form granule cells but many change directions and move outward toward their final positions near the surface of the bulb and become periglomerular cells. The dendrites of these cells become targets for the growth cones of olfactory cell axons that form synapses with them (22,23). A required consequence seems to be that this pattern of cell generation and migration relates directly to the specificity of the target receptor(s) that each cell will ultimately express. This process forms the precise and bilaterally symmetrical topological map of future targets for the growth cones extending from olfactory neurons born in the olfactory epithelium to the glomeruli in the olfactory bulb.

Serpentine receptors play a key and proven role as address molecules targeting the glomeruli. It seems important to examine various regions of the brain and embryo to determine where and when olfactory and VNO receptors are expressed. Clearly, it is reasonable to consider molecules expressed throughout the developing embryo.

There are many molecules other than the olfactory and VNO receptors that have been shown to play an important part in cell surface recognition (8). These molecules fulfill many of the addressing functions needed in an area code system by providing the equivalent of the country codes, area codes, regional codes, etc. One such example is O-CAM, one of a large number of cell surface receptors in the immunoglobulin supergene family (24,25). O-CAM is expressed on a subset of olfactory nerve axons that extend from the four zones of the olfactory epithelium to the specific zones of glomeruli in the olfactory bulb. This molecule is expressed on axons originating in three of the four zones of the olfactory epithelium and on one of the two zones from the VNO region. O-CAM thus seems to provide an excellent candidate for an area code molecule coding for geographic regions rather than for a specific

cellular address. It is assumed that other, probably related receptors will be found on zones in which O-CAM is absent and that these will form part of the combinatorial code.

It may be possible to conceive of genetic, molecular and cellular mechanisms capable of accomplishing the assembly of the two thousand or so target sites in the olfactory bulb. As discussed above, neuronal precursor cells migrate considerable distances along stereotyped routes to lay out a precise, bilaterally symmetrical target map in the olfactory bulb. The mechanisms responsible are completely unknown. The only other example of this extraordinary level of migratory specificity is seen in the targeting of the axonal growth cones as they extend to form synapses in the olfactory bulb. In the absence of any good alternative we shall consider the possibility that members of the olfactory and VNO receptors are expressed in the cells that form the target arrays in the olfactory bulb. In this scenario, molecular interactions of these receptors with each other provide the required specificity for both migration and targeting. Cells may interact in such a way as to form the precise topological map of cells expressing target receptors. One intriguing possibility is suggested by the structure of the receptors themselves and by certain interesting patterns in which these structures are arrayed in the target maps. All of these receptors contain seven helical domains that traverse the membrane and arrange themselves so as to form a pocket at the cell surface. Studies have shown that these pockets provide specific sites for binding ligands. These receptors also display extra-cellular loops of varying size that provide additional specificity for interactions (26). Differences in the amino acid sequences within the domains forming the pockets and loops provide the individual specificity for ligand binding. There is speculation that this structure might also provide specificity for homophilic interactions (27).

Let us consider the notion that these combined binding sites provide the required specificity for both homophilic and heterophilic interactions of these receptors. Homophilic

interactions could account for the target specificity known to occur as the olfactory axons seek specific glomeruli in the olfactory bulb. A possible method for the specificity of cell migration and bulb assembly derives from the observation that serpentine receptors with an unusual type of extracellular loop structure cluster together in both the olfactory epithelium and in the target bulb structure (28). Indeed, several studies suggest that glomeruli are arranged with receptors of similar structure displayed on adjacent glomeruli and within a specific region of the olfactory bulb (29). It seems possible that receptors differing only slightly in the amino acid sequence of the binding sites responsible for homophilic interactions could still interact with relatively high affinity. The binding constant difference could serve to guide neighbors to each other. Other adjacent cells could again have receptors with close but lower affinity. In this manner a type of affinity gradient could be established that, at least theoretically, could help explain the relationships maintained among cells as they migrate and assemble the target map in the olfactory bulb. Such a gradient of receptor affinities would also aid the growth cones of olfactory neurons as they seek their targets in the bulb.

The genetic programs are sophisticated enough to generate and maintain one thousand or more cells, each expressing one receptor gene. Elaborate genetic controls must function to maintain the expression of a single, specific serpentine receptor gene in each of the olfactory stem cells and in its daughter olfactory neurons as they continue to be born throughout life. Furthermore, these controls must allow the expression of only one of the two alleles present in each cell (30). The complexity of this genetic problem is very reminiscent of the similar situation seen in the immune system where sophisticated alterations are made in the germline DNA as specific B or T cells are generated. There too only a single allele is expressed in each cell. The altered DNA sequences are replicated for the life of a stem cell thus accounting for

the lineage memory. Genetic switching therefore remains an attractive aspect of the Area Code Hypothesis, particularly for the control of the expression of the serpentine receptors discussed here. Indeed, it is extremely difficult to imagine that a mechanism utilizing only transcription factors et cetera is capable of mimicking the immune system's single-allele expression and stem cell-specific receptor expression.

Genetic Switches Known to Function in Various Organisms: The earliest proven example occurred of developmentally controlled genetic switching occurred in large colonies of Cyanobacter over two billion years ago (31,32). The same types of cyanobacteria exist today and form large colonies identical to those in the fossil record. In this organism, DNA rings are excised from the germline cell's DNA to form somatic cells that can fix nitrogen for the use of the entire colony. There is good reason to believe that this type of genetic switch evolved very early and has been selected for use in numerous subsequent species because of its efficacy as a means of programming the formation of different cell lineages.

Numerous types of repeats and transposable elements have also been shown to play a role in chromosomal programs, wherein germline DNA is altered as specific cell types are formed. Ciliates, for example, use transposes to excise specific transposon-like elements from germline DNA as a part of the mechanism used to form the somatic macronucleus from the germline micronucleus (33,34). Excision of specific transposable elements occurs in Drosophila as polytene chromosomes are formed from the germline (A. Gould & W. Dreyer, unpub. observations). In another example, it is now known that the telomeres in Drosophila are maintained by two different transposable elements (35). Ribosomal DNA, like telomeres, must be controlled and maintained during development. These chromosomal regions contain numerous tandem copies of rDNA. In *D. melanogaster* specific transposable elements (different from those that maintain telomeres) are associated with rDNA (36). It seems very

possible that they aid in the recombination control required for the maintenance and amplification of these chromosomal regions. Numerous other examples of DNA alterations during development of other organisms can be found in the literature.

The mechanism by which DNA is excised during the development of the immune system is very closely related to many of the examples mentioned above. Indeed, the RAG-1 transposase is evolutionarily related to the enzymes responsible for transposable element rearrangements found in essentially all eukaryotes and even bacterial switches such as the invertrons (37-40). Ten to twenty percent of the DNA of most multicellular organisms is made up of mobile DNA elements, hence large numbers of genes coding for members of the transposase/recombinase family are found in these genomes and according to our hypothesis, some may function in normal development.

The list of confirmed examples of programmed alterations in DNA is now so long that one is quite safe in stating that not all of the repeats and elements that make up a significant part of all chromosomes are "junk DNA." It therefore seems reasonable to examine the possibility that some of the transposon-related elements may play a role in programming the expression of such genes as the serpentine receptors. Again, no other known mechanisms that do not involve alteration of DNA seem adequate to perform the extraordinarily complex programming of gene expression that is discussed here.

One obvious ramification of developmentally programmed DNA alteration is that cells from fully differentiated tissues could not be used to clone new individuals. And in fact this seems to be the case despite the two widely quoted examples of cloning from "differentiated" tissues. Neither the cloning of Dolly from the udder of a sheep (41), nor the cloning of an adult frog from larval frog intestines (42) was proven to have been accomplished from a differentiated cell type. The Dolly experiment has not been repeated

and, even after thirty-six years, no successful repeat of Gurdon's result has been accomplished using confirmed differentiated cells from adult frogs (43). In each case above, the cloned individual was the very rare outcome of numerous experiments, and in both cases an embryonic germ cell could have been the cell actually selected for cloning. This is possible since the sheep which served as a donor for Dolly was pregnant, and since the larval frog intestine is a known site of germ cell migration during development. In contrast to the above reports, the successful use of nuclei derived from blastula cells in the nuclear transplantation experiments pioneered by Briggs and King in 1952 (44) has been reproduced many times and similar procedures have been used by numerous scientists in a variety of species throughout the past forty-six years. Nuclear transplantation from blastulas is compatible with the Area Code Hypothesis because DNA switching has not yet occurred at this stage of development and the cells are therefore totipotential. Thus, in another embodiment, the invention provides a method for obtaining such totipotential germ cells that may have migrated to various tissues (e.g., udder of cows, gonads/testis) and are maintained among the differentiated cells. Such cells are useful as starting material for nuclear transplantation in cloning experiments. In one embodiment, the invention provides a method for producing a specific cell lineage or organ type or an organism comprising obtaining a cell by the method of the invention as described herein. The cell(s) is treated under conditions and for a time sufficient to produce the lineage, organ or organism. For example, methods of producing organisms include nuclear transplantation.

Are repeats and transposon-related elements present in the sequences of the multigene families of serpentine receptors? Figure 2 illustrates one of many examples of the DNA sequences of regions containing genes coding for serpentine receptors. Two serpentine receptors are coded by the DNA sequence illustrated. Note the pattern of elements near both

upstream control regions. We have observed that all known sequences of DNA containing families of serpentine receptors contain sequences related to mobile elements in the non-coding regions. We believe that careful consideration should be given to the possibility that repetitive elements, including some of those illustrated here, may play a role in programming the expression of the very large families of seven-transmembrane receptor genes that have been found.

The data discussed above provide strong support for the notion that such receptors are indeed expressed in numerous tissues other than the olfactory regions. However, the data available at this time do not provide topological details of the expression of these molecules over time and space in the developing embryo. We predict that each receptor will be expressed in a speckled pattern throughout the embryo similar to the locations of the last four digits of phone numbers in geographic locations where they are used repeatedly in combination with other digits to code for different telephone sites. This type of pattern might easily be mistaken for an experimental artifact. A possible example of this may have already been published (14). Monoclonal antibodies developed to fractions of chick embryos correlating to the size of olfactory receptors were used to study expression in chick embryos. Close examination of the expression of olfactory receptors in chick embryos before, during and after notochord formation (see Fig. 6 in ref. 14) reveals numerous such specks not seen in the control. The notochord does indeed express an olfactory receptor but the speckled appearance of other parts of these sections was not noted by the authors. Obviously, more experiments are needed. As one example, the transgenic mice used by Mombaerts et. al. (10) would provide an excellent source of embryos for the study of the expression of olfactory receptors in tissues other than the adult olfactory system illustrated in their publication.

2) Do seven-transmembrane receptors interact with each other as is predicted by the above discussion? We have not yet uncovered any studies bearing directly on this aspect of the hypothesis, but such experiments are feasible. Several of the available excellent methods were used by Yoshihara et al. (24) in their studies of homophilic interactions of O-CAM. We have used an additional method (45). If it can be shown that no homophilic or heterophilic interactions can occur among these receptors other molecules would have to be found to explain the known facts. However, we are not able to offer any reasonable alternative hypotheses at this time.

3) Is there a gradient of closely related receptors on the topological map of glomeruli on the olfactory bulb? While several publications referenced above suggest that this may be true, more work needs to be done. Structural and functional studies of olfactory receptors expressed on neighboring glomeruli are needed to test this notion. Single-cell PCR techniques should facilitate testing of this "receptor gradient" hypothesis.

4) Is the control of the expression of the one thousand or so different serpentine receptors due in part to DNA switches? By now there are so many confirmed examples of the role of DNA alterations in somatic cells of diverse organisms that this part of the hypothesis should be given serious consideration. Several experimental approaches are now capable of providing data relevant to this subject. PCR methods can be used to compare specific stretches of DNA in germ line and somatic cells. DNA libraries from both cell types can also be used to detect specific differences. Protocols are readily available since studies of such differences in cells of the immune system have become commonplace in recent years. We suggest that experiments be carried out to test the notion that the immune system is not alone in the use of mobile-element-related genetic switches in developmental controls of cell lineages.

Our finding that serpentine receptors are expressed in a large number of different tissues has led us to suggest that they may play a central role in coding for cell positioning during embryogenesis. According to this hypothesis, these and other less-specific receptors are used in a combinatorial strategy that provides molecular codes to cell surfaces. Cells use these cell surface codes to guide their assembly of complex three-dimensional structures. The genetic control mechanisms required for the control of these codes are so sophisticated that we suggest they utilize genetic switches related to mobile elements to aid in the control of the expression of codes on embryonic cells. Recombinases from the very large family encoded by mobile elements are candidates for a role in such DNA alterations. Rag-1, a member of this large recombinase family, plays a key role in the genetic events that use mobile element-related switches during the development of the immune system (37,38). A homeodomain that is also found on some of these recombinases (including Rag-1) raises more intriguing questions (39,40).

Table 1. Search of dbEST for olfactory receptor expression in tissues

Tissue-Source of RNA	dbEST ID
Liver/spleen	363864, 354217, 156098, 347337, 363860, 474480, 380307, 153228, 352182, 168933, 141689, 395357, 597381, 510649, 347787, 91639, 18454
Lung	954216, 498479, 498483, 498479, 939339, 1130457, 117342, 113457, 723746, 1644990, 1663526, 1644990, 1574236, 1522628, 1574236, 1522628
Heart	491132, 887378, 461677, 733579, 887598, 235252, 588679, 114840
Placenta	196106, 193508, 206183, 197975, 432198, 1576745
Brain	233020, 283714, 188843, 331341, 42598, 64251, 151626, 712287, 306632
Pineal Gland	1426549, 345459, 1451432
Prostate	939790, 954402, 1670434, 1147520, 939790, 1218857, 868944, 869330, 1218648, 1152521, 1298671, 1298734, 2394196
Testis	1104389, 1414182, 1306625, 1425879, 1318566, 1039952, 1306468, 1326543, 1307226, 1478699, 1534360, 1537251
Breast	300904, 300902, 286374
Mouse mammary	1102549, 1516567
Pregnant uterus	659986, 660104, 790661, 660104
Lung	117342, 1130457, 528502, 723746
Mouse skin	1484894
Retina	937202
Testis	1481318566, 1478698, 1534360, 1537251, 1039952, 1306468, 1587477, 1480933, 1104389, 14780482, 1478698, 1587477
Colon	1305576, 1320848, 1266055, 1224456, 1158847, 1654512, 1669804, 1644878, 1220796, 1604549, 1573063
Ovary	1551728, 1559748, 1551267, 1551288, 1100065, 1551728, 1071934, 1122687, 1548667, 1551267, 1551288
Embryo	938609, 1388065, 16700
Kidney	1698409, 1599751, 1645467, 1669570, 1555887, 1558941, 1599751, 1555887, 1645467

REFERENCES

1. Hood, L., Huang, H. V. & Dreyer, W. J. (1977) J. Supramol. Stn~ct. 7, 531-559.

2. Dreyer, W. J. and Bennett, J. C. (1965). The molecular basis of antibody formation: A paradox. Proc. Nat. Acad. Sci. USA 54, 864.

3. Dreyer, W. J. (1984) in The Impact of Protein Chemistry on the Biomedical Sciences, eds. Schechter, A. N., Dan, A. & Goldberger, R. F. (Academic Press, New York).

4. Dreyer, W. & Roman, J. M. (1984) in Advances in Experimental Medicine and Biology. Gene Expression and Cell-Cell Interactions in the Developing Nervous System, eds.Lauder, J. M. & Nelson, P., Vol. 181 (Plenum Press, New York), pp. 87-97.

5. Kayyem, J. F., Roman, J. M., Von Boxberg, Y., Schwarz, U. & Dreyer, W. J.(1992) Eur. J. Biochem. 208, 1-8.

6. Kayyem, J. F., Roman, J. M., de la Rosa, E. J., Schwarz, U. & Dreyer, W. J.(1992) J. Cell Biol. 118, 1259-1270.

7. Vielmetter, J., Kayyem, J. F., Roman, J. M. & Dreyer, W. J. (1994) J. Cell Biol. 127, 2009-2020.

8. Molecular bases of axonal growth and pathfinding (1997) in Cell ~ Tissue Research, eds. Drescher, U., Klein, R., Sthrmer, C., Faissner, A. & Rathjen, F. G., Vol. 29 (Springer-Verlag, Berlin), pp. 187-470.

9. Mombaerts, P. (1996) Curr. Opin. Neurobiol. 6, 481-6.

10. Mombaerts, P., Wang, F., Dulac, C., Chao, S. K., Nemes, A., Mendelsohn, M., Edmondson, J. & Axel, R. (1996) Cell 87, 675-86.

11. Ressler, K. J., Sullivan, S. L. & Buck, L. B. (1994) Cell 79,1245-55.

12. Vassar, R., Nagi, J. & Axel, R. (1994) Cell 74, 309-318.

13. Drutel, G., Arrang, J. M., Diaz, J., Wisnewsky, C., Schwartz, K. & Schwartz, J. C. (1995) Receptor Channels 3, 33-40.

14. Nef, S. & Nef, P. (1997) Proc. Natl. Acad. Sci. USA 94, 4766-71.

15. Vanderhaeghen, P., Schurmans, S., Vassart, G. & Parmentier, M. (1997)Biochem. Biophys. Res. Commun. 237, 283-7.

16. Bargmann, C. I. (1997) Cell 90, 585-7.

17. Dulac, C. (1997) Neuron 19, 477-80.

18. Dulac, C. & Axel, R. (1995) Cell 83, 195-206.

19. Herrada, G. & Dulac, C. (1997) Cell 90, 763-73.

20. Matsunami, H. & Buck, L. B. (1997) Cell 90, 775-84.

21. Lois, C., Garcia-Verdugo, J. M. & Alvarez-Buylla, A. (1996) Science 271, 978-81.

22. Luskin, M. B. (1993) Neuron 11, 173-89.

23. Klenoff, J. R. & Greer, C. A. (1998) J. Comp. Neurol. 390, 256-267.

24. Yoshihara, Y., Kawasaki, M., Tamada, A., Fujita, H., Hayashi, H., Kagamiyama, H. & Mori, K. (1997) J. Neurosci. 17, 5830-42.

25. Yoshihara, Y. & Mori, K. (1997) Cell Tissue Res. 290, 457-463.

26. Shepherd, G. M, Singer, M. S. & Greer, C. A. (1996) The Neuroscientist 2, 262-271.

27. Singer, M. S., Shepherd, G. M. & Greer, C. A. (1995) Nature 337, 19-20.

28. Kubick, S., Strotmann, J., Andreini, I. & 8reer, H. (1997) J. Neurochem. 69, 465-75.

29. Friedrich, R. W. & Korsching, S. I. (1997) Neuron 18, 737-752.

30. Chess, A., Simon, I., Cedar, H. & Axel, R. (1994) Cell 78, 823-34.

31. Carrasco, C. D. & Golden, J. W. (1995) Microbiology 141, 2479-2487.

32. Haselkorn, R. (1992) Annul Rev. Genet. 26,113-130.

33. Williams, K., Doak, T. G. & Herrick, G. (1993) EMBO J. 12, 4593-4601.

34. Jacobs, M. E. & Klobutcher, L. A. (1996) J Euk. Microbiol. 43, 442-452.

35. Pardue, M. L., Danilevskaya, O. N., Traverse, K. L. & Lowenhaupt, T. K. (1997) Genetica 100, 73-84.

36. Eickbush, T. H., Burke, W. D., Eickbush, D. G. & Lathe, W. C., III (1997) Genetica 100, 49-61.

37. Xu, W., Rould, M. A., Jun. S., Desplan, C. & Pabo, C. O. (1995) Cell 80, 639650.

38. Ramsden, D. A., van Gent, D. C. & Gellert, M. (1997) Curr. Opin. Immunol. 9, 114-120.

39. Spanopoulou, E., Zaitseva, F., Wang, F.-H., Santagata, S., Baltimore, D. & Panayotou, G. (1966) Cell 87, 263-276.

40. Pietrokovski, S. & Henikoff, S. (1997) Mol. Gen. Genet. 254, 689-695.

41. Wilmut, I., Schnieke, A. E., McWhir, J., Kind, A. G. & Campbell, K. H. S. (1997) Viable offspring derived from fetal and adult mammalian cells. Nature 385, 810-813.

42. Gurdon, J. B. (1962) Dev. Biol. 4, 256-273.

43. Cloning: Nuclear Transplantation in Amphibia (1978), ed. McKinnell, R.(U. Minnesota Press, Minneapolis).

44. Briggs, R. & King, T. J. (1952) Proc. Natl. Acad. Sci. USA 38, 455463.

45. Suter, D. M., Pollerberg, G. E., Buchstaller, A., Giger, R. J., Dreyer, W. J. & Sonderegger, P. (1995) J. Cell Biol. 131, 1067-1081.

46. Glusman, G., Clifton, S., Roe, B. & Lancet, D. (1996) Genomics 37, 147-60.

47. Ressler, K. J., Sullivan, S. L. & Buck, L. B. (1993) Cell 73, 597-609.

48. Scott, J. W., Shannon, D. E., Charpentier, J., Davis, L. M. & Kaplan, C. (1997) J. Neurophysiol. 77, 1950-62.

49. Strotmann, J., Konzelmann, S. & Breer, H. (1996) Cell Tissue Res. 284, 34754.

50. Juilfs, D. M., Fulle, H. J., Zhao, A. Z., Houslay, M. D., Garbers, D. L. & Beavo, J. A. (1997) Proc. Natl. Acad. Sci. USA 94, 3388-95

While the invention has been described in detail with reference to certain preferred embodiments thereof, it will be understood that modifications and variations are within the spirit and scope of that which is described and claimed.

WHAT IS CLAIMED IS:

1. A method of obtaining a composition substantially enriched in a specific cell type comprising:

contacting a sample of cells with at least one binding agent specific for an serpentine cell surface marker such that the binding agent binds specifically to a cell or cells having the marker in the sample; and

separating the cell or cells bound by the binding agent from the sample, thereby obtaining a composition substantially enriched in a specific cell type.

2. The method according to claim 1, further comprising separating the cell or cells bound by the binding agent by selecting for at least one additional marker associated with a specific cell type.

3. The method according to claim 2 wherein the additional marker is selected from the group consisting of CD-34, Thy-1, rho, Cdw109, protocadherins and cell adhesion molecules (CAMs).

4. The method of claim 1, wherein the binding agent is selected from the group consisting of a ligand and an antibody.

5. The method of claim 4, wherein the antibody is monoclonal or polyclonal or derivative thereof.

6. The method of claim 1, wherein the binding agent is immobilized on a solid support.

7. The method of claim 1, further comprising analyzing the DNA of the cells.

8. The method of claim 7, wherein the analyzing is on a microchip.

9. The method of claim 7, wherein the analyzing is by Southern Blot analysis.

10. The method of claim 1, wherein the at least one binding agent is multiplexed such that more than one binding agent is utilized simultaneously.

11. A method of identifying a cell type comprising detecting expression of at least one serpentine receptor, wherein the presence of the serpentine receptor is indicative of a cell type.

12. The method of claim 11, wherein the receptor is detected using a binding agent that binds to the serpentine receptor.

13. The method of claim 12, wherein the binding agent is an antibody or derivative thereof.

14. The method of claim 13, wherein the antibody is monoclonal.

15. The method of claim 11, wherein identification is by detecting a polynucleotide encoding the at least one serpentine receptor.

16. The method of claim 14, wherein the polynucleotide is DNA or RNA.

17. The method of claim 12, wherein the binding agent is a ligand.

18. The method of claim 11, wherein the cell is a totipotential germ cell.

19. The method of claim 11 or 15, wherein the detecting is on a microchip.

20. The method of claim 16, wherein the analyzing is by Southern Blot analysis.

21. The method of claim 11, wherein the detetion of the polynucleotide is multiplexed such that more than one probe which binds to the polynucleotide is utilized simultaneously.

22. A cell obtained by the method of claim 1.

23. A method for producing a specific cell lineage or organ type or an organism comprising obtaining a cell by the method of claim 1 and treating the cell under conditions and for a time sufficient to produce the lineage, organ or organism.

24. The method of claim 23, wherein the conditions include nuclear transplantation.

25. A method for delivery of an agent to a specific cell type comprising performing multiplex binding of a ligand or agent to a cell surface receptor, such that delivery is only to a particular cell type.

26. The method of claim 25, wherein at least one of the ligands or agents that binds to the receptor includes a targeting sequence for delivery to the nucleus of the cell.

27. A method of treating a cell proliferative disorder in a subject, the method comprising:
administering an effective amount of an antibody coupled to a first factor wherein the antibody binds to a first serpentine antigen on the surface of a cell associated with the disorder;
administering a second antibody coupled to a second factor, wherein the antibody binds to a second ligand on the surface of a cell, different from the first ligand; and
wherein the first and second factors react to form a cytotoxic drug, thereby inhibiting the growth of the cell.

28. The method of claim 27, wherein the antibody is selected from the group consisting of polyclonal, monoclonal and chimeric antibodies or derivatives thereof.

29. The method of claim 27, wherein the first factor is a prodrug and the second factor is an enzyme that cleaves the prodrug.

30. The method of claim 27, wherein the first factor is a proenzyme and the second factor is an activator of the proenzyme.

31. The method of claim 30, wherein the enzyme is selected from the group consisting of alkaline phosphatases, proteases, arylsuflatses, beta lactamases, penicilin amidases, D-alanyl carboxypeptidases, and cytosine deaminases.

32. The method of claim 27, wherein the adminstering is ex vivo.

33. The method of claim 27, wherein the adminstering is in vivo.

34. A method of detecting a cell proliferative disorder in a sample, the method comprising contacting the sample with a first antibody coupled to a factor wherein the antibody is reactive with an serpentine ligand on the surface of a cell having a cell proliferative disorder; contacting the sample with a second antibody coupled to a second factor, wherein the antibody is reactive to a second ligand on the surface of a cell; and wherein the first and second factors react to form a reporter agent.

35. The method of claim 34, wherein the antibody is selected from the group consisting of polyclonal, monoclonal and chimeric antibodies.

36. The method of claim 34, wherein the first factor is a pro-factor and the second factor is an enzyme that cleaves the pro-factor.

37. The method of claim 34, wherein the enzyme is selected from the group consisting of alkaline phosphatases, proteases, arylsulfatases, beta lactamases, penicilin amidases, D-alanyl carboxypeptidases, and cytosine deaminases.

38. The method of claim 34, wherein the reporter agent is selected from the group consisting of a bioluminescent compound, a chemiluminescent compound, a metal chelating agent, and an enzyme.

39. A method for detecting at least one variation in at least one serpentine polynucleotide in a cell, wherein the variation is indicative of a particular lineage comprising: a) contacting a sample containing nucleic acid isolated from a first cell with at least one probe that hybrides to the at least one serpentine polynucleotide in the sample, wherein the serpentine polynucleotide is associated with a particular lineage or cell type, thereby forming a hybridization complex; and b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of a serpentine polynucleotide having a nucleic acid sequence complementary to the probe, and wherein the presence of the complex is indicative of a particular lineage or cell type.

40. The method of claim 39, wherein the nucleic acid material is amplified by the polymerase chain reaction prior to the hybridizing step.

41. A method for producing a lineage-specific cell type comprising contacting a cell of a first lineage with a lineage determining effective amount of an agent for a time and

under conditions such that the cell is committed to a second lineage, thereby producing a lineage-specific cell type.

42. The method of claim 41, wherein the agent is selected from the group consisting of an antibody, a ligand, a hormone, a growth factor, an antisense oligonucleotide and any combination thereof.

43. The method of claim 41, further comprising isolating cells of the second lineage from the other cells.

44. The method of claim 41, wherein the lineage is effected by contacting the cells with an agent that modulates Notch or related family members.

45. The method of claim 41, wherein the lineage is effected by contacting the cells with an agent that modulates Delta or related family members.

46. A method for maintaining a lineage-specific cell type comprising contacting the cell with a lineage maintaining effective amount of an agent such that the cell is committed to the lineage.

47. The method of claim 46, wherein the agent is selected from the group consisting of an antibody, a ligand, a hormone, a growth factor, an antisense oligonucleotide and any combination thereof.

48. The method of claim 46, wherein the maintaining the lineage is effected by contacting the cells with an agent that modulates Notch or related family members.

49. The method of claim 46, wherein the maintaining the lineage is effected by contacting the cells with an agent that modulates Delta or related family members.

ABSTRACT

The present invention provides methods for identifying and for producing cells of different lineages based on their cell surface display. In particular, the invention provides methods for identifying different lineages based on the cell surface display of serpentine receptors. Methods of producing cells of specific lineages are also disclosed, as are methods of use for such cells, including using totipotential germ cells identified by the methods of the invention for producing tissues, organs or whole organisms.

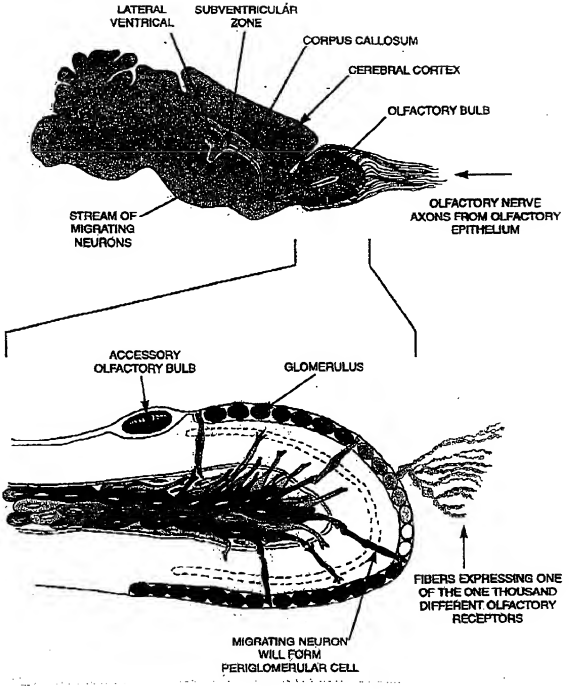
LATERAL VENTRICAL
SUBVENTRICULÁR ZONE
CORPUS CALLOSUM
CEREBRAL CORTEX
OLFACTORY BULB
STREAM OF MIGRATING NEURONS
OLFACTORY NERVE AXONS FROM OLFACTORY EPITHELIUM
ACCESSORY OLFACTORY BULB
GLOMERULUS
FIBERS EXPRESSING ONE OF THE ONE THOUSAND DIFFERENT OLFACTORY RECEPTORS
MIGRATING NEURON WILL FORM PERIGLOMERULAR CELL

FIGURE 1

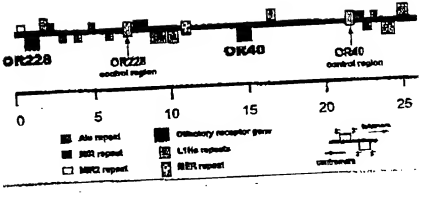
OR228
OR228
control region
OR40
OR40
control region
0
5
10
15
20
25
Alu repeat
MIR repeat
MIR3 repeat
Olfactory receptor gene
L1Hs repeats
MER repeat
between
upstream

FIGURE 2

DECLARATION FOR PATENT APPLICATION

As a below-named inventor, I hereby declare that:

My residence, post office address and citizenship is as stated below next to my name.

I believe that I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled METHOD OF DETERMINING CELL OR TISSUE TYPE BY TRANSMEMBRANE RECEPTOR IDENTIFIERS, the specification of which

__X__ is attached hereto.

____ was filed on _____, 1999 (Attorney Docket No. _____)
as U.S. Application Serial No. _____
and was amended on _____
if applicable (the "Application").

I hereby authorize and request insertion of the application serial number of the Application when officially known.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose to the United States Patent and Trademark Office all information known to me to be material to patentability of the subject matter of the Application as defined in Title 37, Code of Federal Regulations ("C.F.R."), § 1.56.

With respect to the Application, I hereby claim the benefit under 35 U.S.C. Section 119(e) of any United States provisional application(s) listed below:

60/095,148 August 3, 1998
(Application Serial No.) (Filing Date)

With respect to the Application, I hereby claim the benefit under 35 U.S.C. Section 120 of any United States application(s), or Section 365(c) of any PCT International application designating the United States, listed below and, insofar as the subject matter of each of the claims of the application is not disclosed in the prior United States or PCT International application in the manner provided by the first paragraph of 35 U.S.C. Section 112, I acknowledge the duty to disclose to the United States Patent and Trademark Office all information known to me to be material to patentability of the subject matter of the Application as defined in Title 37, C.F.R., Section 1.56 which became available between the filing date of the prior application and the national or PCT International filing date of the Application:

(Application Serial No.)	(Filing Date)	(Status) (patented, pending, abandoned)

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Full name of first inventor: William J. Dreyer

Inventor's signature: _____

Date: _____

Residence: 1875 Devon Road; Pasadena, CA 91103

Citizenship: USA

Post Office Address: 1875 Devon Road
Pasadena, CA 91103

6140839_1.DOC